DOCUMENT RESUME

ED 439 157 TM 030 694

AUTHOR	Papanastasiou, Elena C.
TITLE	Teacher Evaluation: Theories and Practice.
PUB DATE	1999-11-00
NOTE	27p.
PUB TYPE	Information Analyses (070)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Academic Achievement; *Educational Practices; Elementary Secondary Education; Evaluation Methods; *Evaluation Problems; Formative Evaluation; Teacher Effectiveness; *Teacher Evaluation; Theories

ABSTRACT

This paper reviews literature on teacher effectiveness and evaluation. In order for teacher evaluation to be effective in terms of improving teacher performance, it has to be formative. Evaluations also can contribute to the improvement of the performance of the school overall. Teacher evaluation also has effects on the environment outside the school organization, such as the community of parents. Many problems are associated with teacher evaluation, whether by principals, students, or teacher peers. The use of student achievement data for teacher evaluation has also been a source of debate in teacher evaluation literature. To reduce many of the disadvantages associated with evaluation approaches, a combination of methods should be used to assess teacher performance. In addition, teacher evaluations should be performed frequently to capture a realistic and representative picture of each teacher. (Contains 42 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 439 157

Teacher Evaluation- Theories and Practice

Elena C. Papanastasiou

Michigan State University

November, 1999

TM030694

E. Papanastasiou

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

INTRODUCTION

> In all societies, throughout human history, people have educated their children. Indeed, one of the fundamental characteristics of human civilization is a concern for the preparation of the next generation. From one generation to the next, we seek to pass on what we know and have learned, hoping to ensure not merely the survival of our offspring, but that of our culture as well" (Reagan, 1996, p.ix).

Due to this importance of education today, researchers have been looking at how nations are restructuring their funding, governance, teaching and evaluation to determine what effects they have on student learning (Stedman, 1994). The reason is because it is believed that international comparative studies can help researchers and educators identify the factors that promote educational achievement, and which correlate with excellence in school performance (Bradburb & Gilford, 1990; Griffith & Medrich, 1992; Lazer, June 1992). International studies seek to "provide rich collections of data, sophisticated measures of student attainment, contextualized interpretation of the results, content-sensitive scaling without misleading aggregations of different items, and guidance of all data collection and interpretation by precisely stated models of education indicators and their relationships " (Schmidt & McKnight, 1995, p.337). In other words, they examine educational outcomes in the light of different philosophies, and in relation to national characteristics such as social, economic and political systems (Tsayang, 1990). International perspectives always provide better understanding of what works in schools and what doesn't (Stedman, 1994).

Therefore, countries participate in studies such as the Third International Mathematics and Science Study (TIMSS) in order to compare themselves normatively

with other countries and other educational systems. According to Bracey (1996) in the case of the US, the amount of publicity that has been drawn to such comparisons is so large, that the trends on the US domestic achievement tests are no longer relevant in the United States, since the salient reference groups for many nations are students from other countries.

However, when administrators, policy makers or the public are not satisfied with their country's results on such international studies, they try to find scapegoats to blame for these results. In the US, for example, people have tended to frequently focus on blaming teachers for the low achievement of students (Wilson & Wood, 1996). Part of the reason for which teachers are blamed, is because incompetent teachers are currently teaching in US classrooms because the system has failed multiple times to perform the task of a gatekeeper (Mehrens, 1991). Consequently, it is assumed that if we were able to attract more high quality teaching candidates into teacher education programs in the US, student achievement would improve as well.

This comes in contrast with Cyprus, however, where teacher education programs are able to attract the highest quality candidates, as far as their school achievement is concerned (Papanastasiou, 1996). The attractiveness of this profession might be because teachers are guaranteed teaching jobs with high salaries the year after they graduate form the university (Papanastasiou, 1996). However, the student achievement in Cyprus (in terms of the TIMSS results) continues to be low (Mullis, Martin, Beaton, Gonzales, & Kelly, 1997). Therefore, some of the blame for the low achievement in Cyprus is also placed on the teachers again. The reason is because it is assumed that the teachers are not motivated to perform their job well since their motives for entering the profession, are extrinsic (e.g. job security, high salaries) (Papanastasiou, 1996). This

example clearly shows that high achievement levels in school, in terms of the teachers, does not necessarily correspond to quality education.

Therefore, one cannot look at a single variable in order to correct all educational problems within a country. Education is a very complex topic. Part of the reason for its complexity is because we are all complex and unpredictable individuals. We are not robots or computers, where a single input will always produce the same output. On a daily basis, people are exposed to thousands of stimuli which interact with their genetic buildup (Sternberg, 1990). Consequently, not everyone is affected by each stimulus in the same way. The same is the case with education.

Educational systems are also complex systems, whose quality depends on multiple components, such as teachers, students, parents, administrators, teacher training institutions, the school environment (e.g. school SES and school resources), each student's environment, and their opportunities to grow in that environment. However, many critics tend to focus on blaming teachers for the low student achievement scores.

There is no doubt that teachers do have a great impact on student achievement. In a study conducted by Wright, Horn, & Sanders (1997), it was shown that the most important factor that affected student achievement was the teacher. However, as described by Sweeney (1994) , teaching is "an incredibly complex activity requiring hundreds of highly subjective teaching decisions during a hectic day. The instructional decision-making process is confounded by dozens of rapidly changing interacting contextual factors and the need to employ highly developed interpersonal skills to implement instructional decisions" (p.224). Therefore, it is important that efforts are made to ensure the quality of performance throughout a teacher's teaching career. This can be done through various evaluation techniques.

DEFINING QUALITY

Reasonable questions that might be asked in terms of teacher evaluation are, "What is 'good education'?" or "How can quality education be defined?" These are important questions since their answer might differ greatly from person to person and from culture to culture. Is quality education the ability for students to perform well on standardized tests, or is quality education the balanced attempt to cultivate the student's mind and body just like the ancient Greek model? In the case of Cyprus, the twelfth grade students that were in the advanced math and science sections had performed exceptionally well on the TIMSS advanced math and science test. Part of the reason for this exceptional performance might be the fact that the students study very hard during their senior year because they have to prepare for the high stakes and competitive university entrance examinations. Therefore, the students are under a lot of stress, they give up any extracurricular activities, they take private afternoon lessons to prepare them for those examinations, and basically only focus on studying. The result of all this preparation is that the students learn the materials exceptionally well, and might not even have to take any introductory calculus, physics or chemistry classes during their first two years of study in an American college or university. However, is that what we really want from students? Is that quality education?

On the other hand, some people view the American educational system as the other extreme. By the time the students go to high school, they are involved in many extracurricular activities such as athletics, music, theater, public service (volunteer work), etc. Academically, however, many American students do not seem to put a lot of effort in their work. Is this quality education?

From the examples mentioned above it is clear that neither extreme can be considered as a good example of quality education. Where can one draw the line, however, between the mix of hard work for high stakes testing and cultivating a student's

overall body and mind? This question is not easy to answer. Part of the reason for this difficulty is because teaching deals with people who are complex individuals that are not easily predictable. Another reason, is because teaching does not include one single skill that can easily be appraised. Despite the difficulty in defining this construct, 'successful teachers' continues to be an important term for teacher evaluation. In order to evaluate teachers, one has to have a clear understanding of what a successful teacher is like. This construct is also important for the improvement of teacher education institutions. By knowing how effective graduates from various teacher education programs are, the administrators of these programs can make changes to them with sound justifications. Therefore, indicators of teacher quality can concurrently be used as indicators for teacher success. Some indicators of teacher effectiveness/success are defined below. It should be made clear though, that no single teacher attribute or characteristic is adequate to define an effective teacher. As was stated above, multiple sources of information should be collected before any conclusion on a teacher evaluation is reached.

Subject matter knowledge

Many of the current studies dealing with teacher quality have determined that an important characteristic for quality teachers is good subject matter knowledge (Ingressol, 1995). When teachers have a good subject matter knowledge, then they are able to help students understand the core ideas of various topics, create useful cognitive maps, as well as enable the students to connect each topic with everyday life examples and facts (Darling-Hammond, 1999) Despite this fact, in 1991, about a fifth of the US secondary public school students were taught by teachers who did not have at least a minor in subjects such as English, literature, journalism, etc (Ingressoll, 1996). In addition, about a quarter of the US public secondary school students that same year

took math courses by teachers who did not have at least a minor in math or math education. In another study conducted by Ballou and Podgursky, (1995), it was determined that the average cognitive ability of teachers was below the mean when compared to the college educated population. Similar observations were found by Stoddart (1990), who realized that school districts in most states end up employing teachers with emergency credentials to teach courses which they are not certified to teach.

Interpersonal skills

Another characteristic of quality/successful teachers are good interpersonal skills (Valentine, 1992). According to Perlman & McCann (1996), this is very essential since teaching is a communal activity, and teachers have to interact with their students and with other colleagues on a daily basis. Communicating effectively with the students' parents or guardians is also an essential activity since it keeps the parents/guardians informed on their children's' progress (Pelletier, 1995). Such activities can include formal or informal meetings with groups or individuals parents, written communication (e.g. report cards, informal notes), as well as telephone call conferences (Pelletier, 1995).

Professional development

Professional development can be considered as a sign of a quality teacher since it reflects an ongoing effort to improve practice and to keep teachers current (Pelletier, 1995). Such practices can include participation in teacher workshops, special training, additional college courses or advanced degrees, frequent participation in inservice meetings, as well as being a member of teacher organizations, networks, or unions (Pelletier, 1995).

However, it should be made clear that the existence of the variables mentioned above does not guarantee that a teacher will be successful. A deficiency in one area cannot always be compensated by a strength in another area (Mehrens, 1986). For example, if a teacher has all of the characteristics mentioned above, but has very poor classroom management skills, and is unable to control the attention of the students, all the variables mentioned above will be ineffective. This is another complexity that has to be dealt with, in terms of the issue of teacher evaluation.

Theories of teacher education

The way that quality education is defined is extremely important since it also affects the theories on which teacher education programs are based. However, these theories tend to change through the years based on the current research and theories that take place. Some examples of teacher education theories that have appeared in the literature, as described by Tatto (1999) are the following:

1) The behavioristic approach where teacher education is seen as mastering a set of skills to support teachers in the performance of a predetermined task,
2) Humanistic teacher education, which is related to adult development where teachers are encouraged to find their own best ways to function as teachers,
3) Teaching as a craft, where teaching is accumulated by trail and error and is to be found in the wisdom of the experienced practitioners. A master- apprentice relationship is seen as necessary to uncover good practice and to pass it on to the novice,
4) Inquiry oriented teacher education, which gives emphasis to the development of inquiry about teaching and about the contexts in which teaching is carried out. The focus is on helping teachers develop dispositions and skills of critical inquiry to

analyze what they are doing in relation to its influence upon learners, schools and the larger society,

5) Constructivist teacher education argues that teachers need to know that knowledge is constructed by individuals and their society; they need to be able to construct knowledge themselves, within a sociohistorical context in order to help their pupils find meaning in what they learn, and
6) Pragmatist teacher education where truth is modified as discoveries are made and it is relative to the time and place and purpose of inquiry.

These distinctions should not be considered as mutually exclusive. In addition, none of these approaches is universally considered as the best approach to teacher education. These approaches can even vary based on the subject matter that is being taught (Schriven, Wheeler, & Haertel, 1992). However, the theories of teacher education are presented to show that quality education and quality teaching are defined differently based on each approach. For example, in the teacher education approach of teaching as a craft, it might be considered appropriate for a teacher to make mistakes since teaching is believed to be accumulated by trial and error. Making errors might not be acceptable in any other approach, though, since it is generally considered as a sign of bad teaching.

THE NEED FOR TEACHER EVALUATION

Another reason why quality teaching is difficult to define is because teaching can be considered as an art. So like art, there is no clear recipe of ingredients that will always produce the same quality of results when they are combined in a specific way. Researchers do find correlations between indicators of quality teaching with variables such as good subject matter knowledge and highly energetic teachers. However, these

are just correlation, not cause-effect relationships. Therefore, adopting any of these characteristics does not guarantee quality teaching.

Consequently, why should teachers be evaluated if there is no universally accepted definition of a good teacher? A simplistic reason, from my point of view, is because it is easier to determine what a bad teacher is like, and because someone needs to make sure that teachers are performing their roles appropriately. Mehrens (1991), takes a different stance by stressing that the need for evaluating teachers is especially important since colleges and state licensing boards in the US have failed multiple times as gatekeepers. He explains that this is true because a) teacher training institutions tend to admit low quality candidates, b) institutions grant diplomas to graduates that are not competent, c) states license teachers that are not minimally competent, and d) a large proportion of the teacher training institutions have unacceptable teacher preparation programs that have not been shut down (Mehrens, 1991). Therefore, since incompetent teachers have been allowed to enter the workplace, teacher evaluation processes should exist in order to alleviate this problem. Natriello (1990) on the other hand, describes three levels of reasons for which teachers should be evaluated. An individual level effect, an organizational and an environmental level effect.

Teacher level effects / Improvement of teaching skills - Providing feedback

One of the main reasons for teacher testing is to increase teacher quality (Mehrens, 1991; Natriello, 1990; Valentine, 1992). As stated by Andrews (1995) teacher evaluation is "one of the most powerful tools for improving instruction. Effective evaluation by high-quality evaluators will improve teacher-administrator communication, motivate teachers to improve classroom practices, assist in the counseling out of incompetent teachers, and ultimately improve student achievement " (p.73). "Evaluation

could make them better teachers for the benefit of their students, because better teaching might inspire better learning" (Nevo, 1994, p.116).

However, in order for teacher evaluation to be effective in terms of improving teacher performance, it has to be formative, and therefore descriptive rather than judgmental in its nature (Nevo, 1994). This improvement is even more likely to occur if schools or school districts connect their evaluations to professional development courses that are provided in order to foster improvement in teaching practice (Natriello, 1990).

Finally, the practice of teacher evaluation can also provide opportunities to recognize quality teachers (Andrews, 1995). This recognition can serve as a motivation for other teachers to strive for improving their instructional practices. This is especially important for teachers who cannot always be aware of what is expected from them since they do not have many chances of interacting with each other to exchange knowledge and skills (Sweeney, 1994). Therefore, the teacher evaluation process can allow teachers to obtain reassurance about their performance, and feel good about it (Sweeney, 1994).

Organizational level effects

The evaluation process can also effect the improvement of the performance of the school overall (Iwanicki, 1990). Specifically, when evaluations are soundly based, teachers will try to avoid behaviors that caused other teachers to be evaluated negatively, and adopt behaviors that were associated with positive evaluations (Natriello, 1990). This will help develop high standard norms for individual schools.

The organizational level effects of these evaluations can be even greater when the school staff is involved in the definement of the construct of 'good teaching practices'

since teaching expectations will be very clear for the teachers. By allowing teachers to participate in the general evaluation processes (including their own evaluation), teachers will "move from seeing schools as *my* children in my classroom to *our* children in our school and *our* children in our community or society. Teachers [will then become] the source rather than the object of change. " (Peterson & Chenoweth, 1992, p.182).

Environmental level effects

As cited in Haefele (1992), "ineffective teacher evaluation practices have allowed unqualified persons to assume teaching positions, made it difficult to rid education systems of incompetent and unproductive teachers, failed to provide direction for staff development, not adequately recognized outstanding instruction, and failed to provide evidence that will withstand professional and judicial scrutiny (Joint committee on standards for educational evaluation, 1988)" (Haefele, 1992, p.225). This is supported by Tucker (1997), who discusses that although 5-15% of public school teachers are incompetent, less than 1% of those teachers are dismissed.

Therefore, the evaluation process can be effective by justifying practices such as hiring, promoting, or firing teachers (Natriello, 1990; Wilson & Wood, 1996). This information is especially important for the environment outside the school organization, such as the community or the parents that might question various school practices (Natriello, 1990).

Teacher- education (TE) effects

Mehrens (1991) extends his views about testing and supports the use of tests in order to admit students in teacher education programs. The effects of this type of testing is twofold. First, it can increase the level of students that are admitted to such programs,

which will consequently increase the level of instruction. Second, it can also effect teacher education institutions by forcing the institutions that have inadequate levels of admission standards, instructional levels, or exit requirement, to shut down (Mehrens, 1991).

I also agree that teacher education institutions could drastically improve their teacher education programs if proper teacher evaluation processes exist. However, this is not an easy process. Currently, MSU is trying to evaluate its teacher education program. This process if very difficult, however, since a teacher education program is a system with multiple components that all need to be taken into account during the evaluation process.

To be specific, every year, approximately 550 individuals are accepted in the program. These individuals bring with them many personal traits and characteristics. Such traits include their high school achievement, interests, their professional goals, their motivation, as well as other personal characteristics. Once the individuals are selected in the program, they go through a training process that involves taking classes, as well as practical training. Many factors and variables affect the teacher training. Such variables include the students' instructors, their university curriculum, the specific class which they student teach, their cooperative teacher as well as their student teaching supervisor. Eventually, many of these students graduate from the program, and are awarded a Bachelor's diploma.

However, even though the students get a diploma with the completion of their program, no real measure exists to determine the precise link between the inputs, the processes of the teacher education program, and its outcomes. Therefore, more research needs to be done to determine the relationship between some of the program inputs and processes to its outcomes.

The teacher education program at Michigan State University is very broad, with a large amount of inputs, so it has numerous aspects that could be evaluated. An effective way of starting to evaluate a teacher education program is by looking at its direct outcomes – the students that graduate from the program with their skills and knowledge- and try to determine the rate of success that those students have in the workplace. Eventually, this could help the program administrators refine the program with sound justification. Therefore, the main purpose of this evaluation is to determine how successful the graduates of the program are in their careers, or their educational paths that they have taken after their graduation from MSU. It is assumed that successful graduates that obtain teaching positions are also quality teachers. It is possible that if sound teacher evaluation procedures did exist for the state of Michigan, those evaluations could also have guided the improvement the teacher education program at Michigan State University, as well as that of other colleges and universities.

Finally, the quality of education could also be improved if the standards for accepting teachers in the profession were significantly increased. Mehrens (1991) had summarized some concerns that by raising such standards there would be a shortage in the supply of teachers, and especially of minority teachers. Mehrens (1991) continues, however, by explaining that teachers should be employed so that they can teach students effectively- not because minorities or other individuals need to find jobs. These concerns have not been substantiated, however, in the case of Cyprus, where the standards for entering the teaching are very high. On the contrary, these high standards have increased the status of the profession, as well as increased the pool of candidates that wish to enter the teaching profession (Papanastasiou, 1996).

METHODS OF DATA GATHERING

Before attempting to perform any type of teacher evaluation, all activities that will be evaluated need to be defined. "The evaluation process should be constructed on a foundation of valid expectations (criteria)... the committee members should know and understand the 'effective teaching' and 'effective schooling' literature. Because the criteria must be applied through an evaluation system that will produce positive results, the literature on teacher evaluation should also be studied" (Valentine, 1992, p.15).

Regardless of how teaching effectiveness is defined, though, it is well accepted that teaching "involves a complex set of knowledges, abilities, and personal attributes in dynamic interplay, ...[which] cannot be captured by standardized paper-and-pencil tests" (Davey, 1991, p.121). What makes the situation even more complicated, is that there are no technically, logically, educationally, and ethically defensible criteria for good teaching (Dwyer, 1994). Therefore, it is reasonable to expect that teacher evaluation data should be obtained from multiple sources. Such sources can include self-evaluations, peer evaluations, evaluations by principals, students or parents, classroom observation protocols, rating scales, student achievement scores, analysis of instructional materials, student questionnaires, or individual clinical supervisions (Nevo, 1994). The multiplicity of methods is also more fair for minority teachers who do not tend to perform well on standardized tests for reasons such as cultural bias (Davey, 1991).

The multiple forms of evaluation gathering techniques is considered as an advisable practice since it enables the evaluators to get a comprehensive picture of a teacher's effectiveness. Stiggins (1989) agrees with the above by stating that "we must shift the focus of our consideration of achievement data in teacher evaluation away from test results derived from centralized, standardized testing programs and towards results derived from teachers' classroom assessments of student achievement. If we help teachers to use high-quality, classroom-level student achievement information to

determine if the instruction is working, then we take a major step towards helping them tap a data source that can produce information needed to establish important professional developmental goals" (p.10).

The Tennessee Career Ladder program is an example of a program that uses multiple sources of data for their teacher evaluation system. Specifically, they obtain data from student questionnaires, principal questionnaires, student questionnaires, portfolios, dialogues, professional skills' tests, as well as evaluator consensus. The content areas that the data relate to are planning, teaching strategies, evaluation, classroom management, and leadership (Rakow & Mclarty, 1990). Another example is that of the Cave Creek Unified school district in Arizona which proposed a composite score for evaluating teachers that consisted of 10% student feedback, 20% student achievement, 10% PGP (Professional Growth Plan) accomplishments, 40% principals ratings, and 10% peer evaluations (Manatt & Price, 1994).

Problems with teacher evaluation data gathering procedures

"States and school districts have been using a variety of evaluation methods to assess the job performance of teachers on a continuous basis including classroom observations by principals and other administrators, rating scales, evaluation by students, and the use of student achievements. None of these methods is clean of criticism" (Nevo, 1994, p.114). For example, not all teacher evaluation methods are equally valid and reliable.

Wilson and Wood (1996) acknowledge that there are many problems associated with teacher evaluation instruments. They discuss how teacher evaluation instruments are not sensitive to measuring innovative teachers, or differences in teaching techniques and strategies across different content areas. "However, rather than resisting such

evaluations on the basis of their limitations, teachers could help develop a more realistic perspective on accountability, accept the need of being accountable to their schools, to the parents of their students, and to their school districts, and work together with their school districts in designing appropriate systems for teacher evaluation" (Nevo, 1994, p.114).

One type of problem associated with teacher evaluation, as discussed by Frase and Streshly (1994) discusses how teacher evaluations tend to be grossly inflated due to tenure laws and other union contract regulations. In a study of six eastern school districts in the US that was conducted by the authors, it was found that teacher evaluation ratings were positively skewed towards the high side of the rating scale. In most districts, none of the teachers (not even the probationary teachers) were rated as below standard. This is especially alarming since ratings obtained from observations of the teachers in these highly rated school districts showed very poor instructional practices such as drill and practice exercises, students asleep or not being attentive in class, or classes unattended by teachers.

Problems with student evaluations

A problem with students' evaluations of their teachers, is that no evidence exists that it actually leads to improved teaching (Andrews, 1995). As explained by Andrews (1995) "there are no face-to-face meetings to discuss the findings with students. The student evaluators remain anonymous and there is no follow-up conducted if student concerns are found. In addition, there is no documented evidence to show that student ratings lead to administration placing faculty members into remediation or removing poor faculty. It appears that poor instruction is not identified through use of student ratings" (Andrews, 1995, p.31).

Problems with peer evaluation

A problem with peer evaluations as discussed in Andrews (1995) by Centra (1979) is that "'colleague evaluation' were even more generous than students' evaluations of their teachers. Some 94 percent of the peers judged the teaching of those being evaluated as 'excellent' or 'good'. He also found them to be statistically unreliable. The highest correlation he could find in the research among ratings by different colleagues was .26 for each item. He concluded that 'this low reliability casts doubts on the value of colleague ratings as they were collected in the study' (p.75)" (p.27). However, I believe that a possible cause of this low reliability and the low correlations might be due to the small variance of the scores that were actually assigned to the teachers by their peers.

Problems with principal's evaluations

"Another concern, noted by Scriven, (1987), is the questionable ability of the principal to evaluate teachers of subject areas foreign to the principal's background, or to evaluate teachers of new curricular programs" (Haefele, 1992, p.338)."One explanation of principals' hesitation to evaluate teachers rigorously may be a lack of confidence in the evaluator role. Principals seem unable and, in many cases, unwilling to devote sufficient time to do the evaluation task adequately" (Haefele, 1992, p.338). Another reason, is that principals are rarely held accountable for their teacher evaluation practices (Bridges, 1990). This is consistent with Wilson and Wood(1996) who discuss how unless administrators are trained in observation and in the use of a particular instrument, the element of rater subjectivity stemming from rater judgement will always be present.

Student achievement data

The use of student achievement data have been a source of debate in the teacher evaluation literature. On the one hand, it appears reasonable that student achievement data provide concrete evidence on the effectiveness of each teacher. The assumption behind this practice is that the lower the student scores, the poorer the teacher performance (Conley & Bacharach, 1990). Another supporter of the use of student achievement data is Stiggins (1989) who believes that "if teachers are held accountable for student achievement, then teacher and student performance will improve" (p.3).

On the other hand, other researchers believe that this is an unacceptable indicator of teacher performance since multiple factors besides teachers account for the variation in student achievement. Such factors are the students' academic abilities, the district's curriculum, and the workload that is assigned to individual teachers each year (Conley & Bacharach, 1990). One should also keep in mind that the performance of students in a particular grade level does not rely solely on the teaching that has taken place in that same year. A student's performance "is something that students acquire over many years, as they pass from classroom to classroom, grade to grade, and building to building. Teachers in junior high/middle school, for example, are dependent on the educational experiences that teachers in elementary schools provide students. Teachers are also dependent on each other to maintain consistent policies, such as homework and discipline practices from classroom to classroom" (Conley & Bacharach, 1990, p.314).

In addition, it is not uncommon for principals to assign the most challenging students to the most efficient teachers. Consequently, student achievement scores

cannot be useful in providing useful information to teachers and administrators about the specific abilities of teachers. In addition, student achievement data cannot provide any useful and constructive information on what teachers should improve or continue doing while teaching. "This failure often results in teachers believing that they can do very little to control the results of their evaluation. In turn, this lack of efficacy will often result in teacher indifference and possible burnout" (Conley & Bacharach, 1990, p.311). Therefore, Wilson and Wood (1996) provide a middle solution suggest that it is more appropriate if teacher evaluations are made based on student progress rather than at looking at absolute achievement scores.

INVOLVEMENT OF TEACHERS IN THE EVALUATION PROCESS

"Teacher evaluation is usually perceived as a means to control teachers, to motivate them, to hold them accountable for their services, or to get rid of them when their performance is poor. Thus, teacher evaluation has the image of something that was invented *against* teachers rather than *for* teachers" (Nevo, 1994, p.109). This might be due to the fact that teachers do not understand the evaluation process, and are not always made aware of how they will be evaluated and how they can benefit from such efforts. According to Nevo(1994), an understanding of the general teacher evaluation process and techniques can help teachers improve their own self-evaluation and teaching performance. "Teachers who understand how teaching is being evaluated could not only improve their self-evaluations; they could also benefit in preparing themselves for being evaluated by others or demonstrating the quality of their skills and performance to designated audience.... Teachers could also learn how to collect and organize evaluative information regarding their competence and teaching experience that would help them to win a teaching job, to use evaluation (feedback to improve their

teaching performance, be accountable to the parents of their students, to negotiate their teaching better evaluation agreements with their school districts, or to get national recognition as outstanding teachers" (Nevo, 1994, p.109). In addition, increased teacher involvement in evaluation can also accommodate a greater variety of teaching styles that cannot be reflected in standardized evaluation methods (Peterson & Chenoweth, 1992).

Valentine (1992)" adds that "for change to be effective and lasting, personnel affected by the change must have input into the development, refinement, and implementation of change. They must understand and support the philosophical basis of the change and be able to envision the positive effects of the change. Participatory development is crucial. Without participatory development, the change, regardless of how appropriate, will not provide the optimum results over a period of years" (Valentine, 1992, p.15). In other words, all personnel must understand the whole evaluation process in order for it to be effective.

Valentine then continues by saying,

"Because teachers and principals are the key actors in the evaluative drama, they must be meaningfully involved in the development, refinement, and implementation of the system. ... Whether the evaluative system is newly developed, recently refined, or unchanged for years, the essentials of the evaluation process should be reviewed with all staff each year. ... Preinforming those affected reinforces the emphasis on fairness and the developmental philosophy (Valentine, 1992, p.34).

Peterson and Chenoweth (1992) have suggested ways with which teachers can get more involved in teacher evaluation. Some of these methods are,

a) gaining control of deciding which types of data sources should be used for their own evaluations

b) peer gathering data where peer teachers assist in the evaluation process

c) review of judgements of data where teachers can serve on panels that make decisions on retention, promotion, etc.

d) district decision making where the teachers can take part in the decision making processes, such as setting standards for quality teacher evaluation judgements.

CONCLUSION

It is well accepted that there are problems associated with teacher evaluations. Some of the problems deal with the unreliability and lack of validity of the methods with which the data were obtained. Other problems deal with some of the negative consequences of assessment. In addition, no single evaluation instrument can capture all aspects of a teachers performance and skills. Despite these disadvantages, though, there are many advantages associated with teacher evaluations that have been described in this paper. In addition, since teachers make such an impact on millions of student's lives, I believe that they should be held accountable for their performance. Therefore, to reduce many of the disadvantages associated with evaluation instruments, I believe that a combination of instruments should be used to assess teachers' performance. In addition, teacher evaluations should be performed very frequently in order to capture a realistic and representative picture of each teacher.

Finally, teachers should realize that they should not be afraid of evaluation. "Research report after research report tells us that educators' motives for being in the profession are altruistic. They want to serve; they want to help others learn. Constructive feedback is fundamental to achieving this goal" (Frase & Streshly, 1994, p.51). Therefore, teacher assessment (as well as the feedback obtained from such assessments) can be considered as necessary prerequisites to assure teacher and teaching quality. People in the medical profession have to pass through various types of

examinations before being able to actually practice the profession. So why should the case with teachers be different?

BIBLIOGRAPHY

Andrews, H., A. (1995). Teachers can be fired! The quest for quality. Chicago: Catfeet press.

Ballou, D., & Podgursky, M. (1995). Recruiting smarter teachers. Journal-of-Human-Resources, 30(2), 326-338.

Bracey, G. W. (1996). International comparisons and the condition of American education. Educational Researcher, 25(1), 5-11.

Bradburb, M. B., & Gilford, D. M. (Eds.). (1990). (1990). A framework for international comparative studies in education. :. Washington DC: National Academy Press.

Bridges, E. M. (1990). Evaluation of tenure and dismissal. In J. Millman & L. Darling-Hammond (Eds.), The new handbook of teacher evaluation. (pp. 147-157). London: Sage publications.

Conley, S. C., & Bacharach, S. B. (1990). Perfromance appraisal in education. A strategic consideration. Journal of personnel evaluation in education, 3, 309-319.

Darling-Hammond, L. (1999). The case for University-based teacher education. In R. A. Roth (Ed.), The role of the University in the preparation of teachers (pp. 13-30). Philadelphia, PA: Palmer Press.

Davey, B. (1991). Evaluating teacher competence through the use of perfromance assessmenttasks: An overview. Journal of personnel evaluation in education, 5, 121-132.

Dwyer, C. A. (1994). Criteria for performance-based teacher assessments: validity, standards, and issues. Journal of personnel evaluation in education, 8(2), 135-150.

Frase, L. E., & Streshly, W. (1994). Lack of accuracy, feedback, and commitment in teacher evaluation. Journal of personnel evaluation in education, 8(1), 47-57.

Griffith, J. E., & Medrich, E. A. (1992). What does the United States want to learn from international comparative studies in education? Prospectus, 22(2), 476-485.

Haefele, D., L. (1992). Evaluating teachers: an alternative model. Journal of personnel evaluation in education, 5, 335-345.

Ingressol, R. M. (1995). Teacher Supply and Demand in the U.S. Paper presented at the Paper presented at the Annual Meeting of the American Statistical Association.

Iwanicki, E., F. (1990). Teacher evaluation for school improvement. In J. Millman & L. Darling-Hammond (Eds.), The new handbook of teacher evaluation. (pp. 158-171). London: Sage publications.

Lazer, S. (June 1992). Learning about the world (22-CAEP-05): The International Assessment of Educational Progress.

Manatt, R. P., & Price, P. P. (1994). Five factor teacher performance evaluation for career ladder placement. Journal of personnel evaluation in education, 8, 239-250.

Mehrens, W. A. (1986). Measurement specialists: motive to achieve or motive to avoid failure? Educational measurement: Issues and practice, 5(4), 5-10.

Mehrens, W. A. (1991). Social issues in teacher testing. Journal of personnel evaluation in education, 4, 317-339.

Mullis, I. V. S., Martin, M. O., Beaton, A. E., Gonzales, E. J., & Kelly, D. L. (1997). Mathematical achievement in the primary school years: IEA's third international mathematics and science study. Boston College. Chestnut Hill, MA: TIMSS international study center.

Natriello, G. (1990). Intended and unintended consequences: Purposes and effects of teacher evaluation. In J. Millman & L. Darling-Hammond (Eds.), The new handbook of teacher evaluation. Assessing elementary and secondary school teachers. (pp. 35-45). London: Sage Publicaions.

Nevo, D. (1994). How can teachers benefit from teacher evaluation? Journal of personnel evaluation in education, 8(2), 109-117.

Papanastasiou, E. (1996). A comparative study of the factors that have influenced students at The Pennsylvania State University and at The University of Cyprus to major in elementary education. Unpublished Honors thesis, The Pennsylvania State University, University Park.

Pelletier, C. M. (1995). A handbook of techniques and strategies for coaching student teachers. Boston:MA: Allyn and Bacon.

Perlman, B., & McCann, L. I. (1996). Recruiting good college faculty. Bolton, MA: Publishing Company.

Peterson, K., D., & Chenoweth, T. (1992). School teachers' control and involvement in their own evaluation. Journal of personnel evaluation in education, 6. 177-190.

Rakow, E., A., & Mclarty, J., R. (1990). Equating a multiple-data-source evaluation system. Journal of personnel evaluation in education, 3, 261-274.

Reagan, T. (1996). Non-Western educational traditions. Alternative approaches to educational thought and practice. Mahwah, NJ: Lawrence Erlaub Associates.

Schmidt, W. H., & McKnight, C. C. (1995). Surveying educational opportunity in mathematics and science: an international perspective. Educational Evaluation, 17(3), 337-353.

Schriven, M., Wheeler, P., & Haertel, G. (1992). Teacher evaluation models project (TEMP). TEMP memo 10. (Vol. TEMP6b,). Kalamazoo, MI.

Scriven, M. (1987). Validity in personnel evaluation. Journal of personnel evaluation in education, 9-23.

Stedman, L. C. (1994). Incomplete explanations: the case of U.S. Performance in the international assessment of education. Educational researcher, 23(7), 24-32.

Sternberg, R. J. (1990). Metaphors of Mind. Conceptions of the Nature of intelligence. Cambridge: Cambridge University Press.

Stiggins, R. J. (1989). A commentary on the role of studentachievement data in the evaluation of teachers. Journal of personnel evaluation in education, 3(1), 7-15.

Stoddart, T. (1990). Los Angeles unified school district intern program: recruiting and preparing teachers for an urban context. Peabody journal of education., 67(3), 84-122.

Sweeney, J. (1994). new paradigms in teacher evaluation: the SBESD model. Journal of personnel evaluation in education, 8(3), 223-237.

Tatto, M. T. (1999, June 28-30, 1999). Conceptualizing and Studying Teacher Education across World Regions: An Overview. Paper presented at the Conference on Teachers in Latin America: New Perspectives on their Development and performance, San Jose, Costa Rica.

Tsayang, G. (1990). Comparative education, an overview. In G. W. Kouwenhoven, S. G. Weeks, & C. Mannathoko (Eds.), Proceedings of the comparative education awareness seminar (pp. 3-12). Botswana: Printing and Publishing Co. Botswana.

Tucker, P. D. (1997). Lake Wobegon: Where all teachers are competent (or have we come to terms with the problem of incompetent teachers?). Journal of personnel evaluation in education, 11, 103-125.

Valentine, J. W. (1992). Principles and practices for effective teacher evaluation. Boston: Allyn and Bacon.

Wilson, B., & Wood, J., A. (1996). Teacher evaluation: A national dilemma. Journal of personnel evaluation in education, 10, 75-82.

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. journal of personnel evaluation in education, 11(1), 57-67.